Phenotype-genotype relationships of *Escherichia coli* O157 and O26 isolates from New Zealand

A thesis presented in partial fulfilment of
the requirements for the degree of
Master of Science
in
Microbiology
at Massey University,
Manawatū, New Zealand.

Tessy Thankachen George

2017

mEpiLab
Institute of Veterinary, Animal and Biomedical Sciences
Massey University

Abstract

Shiga toxin-producing *Escherichia coli* (STEC) is a zoonotic pathogen responsible for causing severe manifestations of gastroenteritis in humans worldwide. STEC is transmitted to humans through the consumption of compromised food and water, or direct animal contact. *Escherichia coli* (*E. coli*) O157 and *E. coli* O26 are considered to be among the important STEC serogroups worldwide due to their association with outbreaks and clinical cases of haemolytic uraemic syndrome and haemorrhagic colitis. This study is concerned with the application of phenotypic microarray technology and high-throughput sequencing technology to characterise *E. coli* O157 and *E. coli* O26 New Zealand isolates.

In this study, 190 phenotypes of carbon sources utilisation were studied in isolates belonging to *E. coli* O157 and O26 serogroups. The isolates were divided into groups based on the fermentation of sorbitol and rhamnose in *E. coli* O157 and O26, respectively. All the isolates respired on approximately 40% of the carbon sources. The sorbitol positive *E. coli* O157 isolates respired on more carbon sources compared to the sorbitol negative *E. coli* O157, rhamnose negative and rhamnose positive *E. coli* O26 isolates.

Genomic analysis showed that *E. coli* O157 isolates had shorter genomes compared to those of the *E. coli* O26 isolates. The core genome comparisons revealed differences between and within the *E. coli* O157 and O26 serogroups. Clustering of *E. coli* O157 and O26 isolates based on sorbitol and rhamnose fermentation, respectively, was observed.

The results obtained from this study illustrated that phenotypic and genotypic differences existed within the *E. coli* O157 isolates. The findings of the current study also demonstrated that while the *E. coli* O26 isolates had similar phenotypic characteristics, genotypic differences existed within the isolates.

Acknowledgements

I sincerely thank my supervisors Dr Anne Midwinter, Associate Professor Patrick Biggs, Dr Julie Collins-Emerson and Professor Nigel French for their valuable suggestions and guidance throughout my study.

Thank you to Massey University and the Institute of Veterinary, Animal and Biomedical Sciences for providing me financial support and funds towards my research. Thank you to Kathy Hamilton, Professor Kevin Stafford and Debbie Hill for their administrative support.

I sincerely thank Dr Ben Phiri for proof-reading my thesis. I would like to thank Dr Jonathan Marshall for his helpful tips with the software R and Simon Verschaffelt for his technical support. Thank you to the staff and postgraduate students at mEpiLab, Massey University, I enjoyed working alongside all of you. I would like to thank Sameer Siddiqui, Rima Shrestha, Saroj Yadav, Punja Prasertsri, Antoine Nohra, José Solis-Ramirez, Kruno Bojanić, Doris Adeyinka and Shirli Notcovich for your kindness and friendship.

Most importantly, I would like to thank my family for their love, support and encouragement.

Let all that I am praise the Lord;
May I never forget the good things He does for me.
Psalm 103:2

Abbreviations

%	Percentage
°C	Degree Celsius
IIA^{glc}	Glucose-specific IIA protein
Acetyl-CoA	Acetyl coenzyme A
ADJ	Adjusted
AE	Attaching and effacing lesion
AFLP	Amplified fragment length polymorphism
argW	Stx-encoding bacteriophage insertion site
ATP	Adenosine triphosphate
BAC	Bacterial artificial chromosome
BD	Becton Dickinson and Company
BGI	Beijing Genomics Institute
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BWA	Burrows-Wheeler Alignment
cAMP	Cyclic adenosine monophosphate
CDC	Centers for Disease Control and Prevention
CO_2	Carbon dioxide
COGs	Clusters of Orthologous Groups of proteins
CRP	cAMP receptor protein
CTA	Cystine trypticase agar
CT-RMAC	Rhamnose MacConkey medium containing cefixime and tellurite
CT-SMAC	Sorbitol MacConkey medium containing cefixime and tellurite
DAEC	Diffusely adhering *E. coli*
DNA	Deoxyribonucleic acid
dsdA	D-serine deaminase or dehydratase
dsdC	LysR-type transcriptional regulator
dsdCXA	D-serine tolerance locus
dsdX	D-serine transporter
EII	Enzymes II
eae	*E. coli* attaching and effacing

EAEC	Enteroaggregative *E. coli*
EAHEC	Enteroaggregative haemorrhagic *E. coli*
EDP	Entner-Doudoroff pathway
EHEC	Enterohaemorrhagic *E. coli*
ehxA	Haemolysin
EIA	Enzyme immunoassay
EIEC	Enteroinvasive *E. coli*
EPEC	Enteropathogenic *E. coli*
ERL	Enteric Reference Laboratory
espA	*E. coli* secreted protein
espB	*E. coli* secreted protein
espD	*E. coli* secreted protein
espP	Protease
ESR	Institute of Environment Science and Research Limited
ETEC	Enterotoxigenic *E. coli*
E. coli	*Escherichia coli*
E. fergusonii	*Escherichia fergusonii*
*fli*C	Flagellin
fliC_{H11}	Flagellin H11
fyuA	High-pathogenicity island gene
g	Grams
Gb	Gigabases
GUD	beta-glucuronidase
H	Flagellar antigen
h	Hours
HC	Haemorrhagic colitis
HGT	Horizontal gene transfer
HPI	High-pathogenicity island
HPr	Heat-stable protein
HUS	Haemolytic uraemic syndrome
irp2	High-pathogenicity island gene
K	Capsular antigen
KEGG	Kyoto Encyclopedia of Genes and Genomes

LEE	Locus of enterocyte effacement
LGT	Lateral gene transfer
M	Molar
Mb	Megabases
MDa	Megadalton
MEE	Multilocus enzyme electrophoresis
µl	Microlitres
ml	Millilitres
Min	Minutes
MLST	Multilocus sequence typing
MPI	Ministry of Primary Industries
NADH	Nicotinamide adenine dinucleotide
NCBI	National Centre for Biotechnology Information
ng	Nanograms
NI	North Island
nle	Non-LEE effector
NMD	National Microbiological Database
NZ	New Zealand
NZGL	New Zealand Genomics Limited
O	Somatic antigen
PAI	Pathogenicity island
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEP	Phosphoenolpyruvate
PFGE	Pulsed field gel electrophoresis
pH	Potential of hydrogen
PM	Phenotypic microarray
PPP	Pentose phosphate pathway
PTS	Phosphotransferase system
QC	Quality control
RFLP	Restriction fragment length polymorphism
RMAC	Rhamnose MacConkey medium
RN	Rhamnose non-fermenting or negative

RNA	Ribonucleic acid
RP	Rhamnose fermenting or positive
SBI	Shiga toxin-encoding bacteriophage insertion
sbcB	Stx-encoding bacteriophage insertion site
SBS	Synthesis by synthesis
SI	South Island
SMAC	Sorbitol MacConkey medium
SN	Sorbitol non-fermenting or negative
SP	Sorbitol fermenting or positive
STEC	Shiga toxin-producing *E. coli*
Stx, *stx*	Shiga toxin
stx1	Shiga toxin 1
stx2	Shiga toxin 2
stx2c	Subtype Shiga toxin 2
T3SS	Type III secretion system
TTC	2,3,5-triphenyl tetrazolium chloride
USA	United States of America
UTI	Urinary tract infection
v	Version
VTEC	Verotoxin producing *E. coli*
wrbA	Stx-encoding bacteriophage insertion site
yehV	Stx-encoding bacteriophage insertion site
yhaJ	LysR-type transcriptional regulator
yhaO	Inner membrane transporter
yhaOMKJ	D-serine sensory locus

Contents

List of Tables

List of Figures